

White-light interferometry using a channeled spectrum. 2. Calibration methods, numerical and experimental results

Chengxing Zhai,* Mark H. Milman, Martin W. Regehr, and Paul K. Best

Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California 91109, USA

*Corresponding author: Chengxing.Zhai@jpl.nasa.gov

Received 9 March 2007; revised 7 September 2007; accepted 13 September 2007;
posted 14 September 2007 (Doc. ID 80885); published 9 November 2007

In the companion paper, [Appl. Opt. **46**, 5853 (2007)] a highly accurate white light interference model was developed from just a few key parameters characterized in terms of various moments of the source and instrument transmission function. We develop and implement the end-to-end process of calibrating these moment parameters together with the differential dispersion of the instrument and applying them to the algorithms developed in the companion paper. The calibration procedure developed herein is based on first obtaining the standard monochromatic parameters at the pixel level: wavenumber, phase, intensity, and visibility parameters via a nonlinear least-squares procedure that exploits the structure of the model. The pixel level parameters are then combined to obtain the required “global” moment and dispersion parameters. The process is applied to both simulated scenarios of astrometric observations and to data from the microarcsecond metrology testbed (MAM), an interferometer testbed that has played a prominent role in the development of this technology. © 2007 Optical Society of America

OCIS codes: 120.3180, 120.5050, 120.5060.

1. Introduction

Precision white-light interferometry is a linchpin technology for ambitious stellar astrometry applications. The Space Interferometry Mission (SIM) Planet Quest mission is a main driver for this technology development with broadband path length delay measurement error requirements of a few tens of picometers [1,2]. These objectives have already been met in the microarcsecond metrology testbed (MAM) using a relatively narrowband signal [3,4]. Extending this performance to broadband applications requires a substantial modification of the existing monochromatic algorithms. The companion paper [5] developed a concise broadband modeling scheme that extends monochromatic models by introducing just a few key parameters.

The key parameters that emerge in this model are all related to the first few moments of the combined source spectral energy distribution and instrument transmission function. The first moment, which is the weighted mean wavenumber across the passband, is

the monochromatic approximation. The higher order moments capture the fringe envelope effects and small phase distortions. In addition, a few parameters characterizing the instrument differential dispersion are also incorporated into this model. Because the estimation algorithm developed in the companion paper is to first order insensitive with respect to first moment error, the accuracy with which this parameter needs to be characterized is relaxed relative to the conventional monochromatic algorithms (e.g., least squares). And since the higher order moments are small corrections to the monochromatic model, they also do not require extremely accurate characterization.

This paper develops and implements the end-to-end process of calibrating these parameters and applying them to the algorithms developed in the companion paper [5]. The process is applied to both simulated astrometric observations and to data from the MAM testbed, which has played a prominent role in the development of the SIM technology program [3,4].

The missing component in this process is the calibration step to estimate the moment parameters and the instrument differential dispersion. The calibra-

tion procedure developed herein is based on first obtaining the standard monochromatic parameters at the pixel level: wavenumber, phase, intensity, and visibility parameters, and then combining them to obtain the “global” moment and dispersion parameters at the spectral channel level, which is obtained by binning the pixels.

A nonlinear least-squares problem is posed to estimate the pixel level parameters. The monochromatic fringe model is linear in the complex visibility and nonlinear only in the wavenumber. This splitting of the parameters can be exploited to reduce the nonlinear estimation of the wavenumber parameter to a simple well-behaved scalar problem. Once the wavenumber estimate is obtained, which is actually the estimate of the weighted mean wavenumber within the pixel (the precise quantity that is sought) the other parameters are estimated with standard monochromatic phase shifting interferometry (PSI) techniques (e.g., least squares).

Alternative methods for pixel level calibration have been investigated based on taking the Fourier transform of the response function [6]. The principal disadvantages to the transform approach stem from the usual difficulties associated with a finite data record, e.g., the resolution depends on the stroke length; the resulting smearing of peaks makes it difficult to determine maxima, the need to interpolate the transformed data, etc. Moreover, accommodations in the transform need to be made for imperfections in the mechanical modulation of the fringe. We remark that a number of hybrid transform methods have been developed for broadband applications to address these issues [7–10]. For our application we have the luxury of stitching together the readily attainable information gathered at the pixel level.

In principle, the procedure developed here is immune to the difficulties discussed earlier. However a sensitivity analysis of the nonlinear approach shows that the wavenumber parameter variance decreases quadratically with the stroke length for a fixed number of photoelectrons. Thus there is a significant SNR advantage for using a longer stroke. The other parameters are shown to be independent of the stroke length. The errors associated with the monochromatic model of the single pixel response are also investigated. Because the diffraction and charge diffusion effects broaden a pixel’s spectral response and thus narrow the coherence envelope, the monochromatic model breaks down for large group delays, which are associated with using long stroke data. Even at the pixel level, neglecting the coherence envelope effect can lead to errors in some of the parameters deleterious at the SIM’s required level of performance. To remedy this situation, a modification to the nonlinear least-squares problem is introduced to remove the coherence envelope effect in the data via a projection operator that annihilates most of the signal contributed by the envelope function.

With the end-to-end procedure in place, a number of simulation and experimental studies were performed. The initial study concerned just the wave-

number calibration procedure. The method was applied to laser interference fringe data from the MAM testbed. The estimated wavelength agreed to within 10 pm of the independently calibrated value. Moreover the observed sensitivity of the error conformed to our analysis of sensitivity to stroke length. End-to-end studies were also conducted on testbed data. Notably it was shown that the broadband algorithms are capable of tracking 100 nm path length delay changes to 50 pm accuracy. A monochromatic algorithm applied to the same data led to errors larger than 1 nm. An important mode for SIM astrometric observation requires “chopping” between stars of a different stellar type. The MAM testbed is not suitable for testing this mode of observing, so to characterize the calibration and fringe estimation procedure a number of simulation studies were conducted using the Star Light Model (SLIM) white-light simulation tool [11]. A new facility is currently being built for this purpose.

2. Overview of White-Light Modeling

In this section, we will briefly review the channeled spectrum approach to white-light fringe estimation developed in the companion paper [5] with particular attention to the model parameter assumptions that were made. A main focus of the present paper is the development of calibration methods that are particularly suited to the overall approach. This is in contrast to the fringe estimation process that simultaneously solves for delay and instrument/source spectral parameters [12].

The paradigm for calibrating the modeling parameters and applying them to the algorithms is shown in Fig. 1. We first use a least-squares calibration algorithm to estimate the pixel fringe parameters, namely, the intensity, visibility, wavenumber, and phase for each pixel based on the measured pixel fringe data. Next we compute the moments of the combined source and instrument transmission function and dispersion parameters required by the white-light algorithms for the spectral channels used in the observation using the calibrated pixel intensities, visibilities, wavenumbers, and phases from the previous step. Finally, with the moments and dispersion parameters computed for the spectral channels used, we apply the white-light algorithms developed

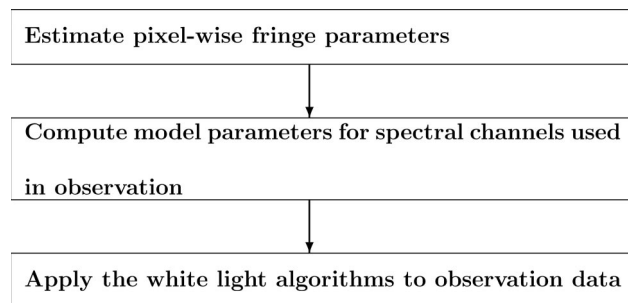


Fig. 1. Flow chart of the end-to-end calibration and estimation procedure.

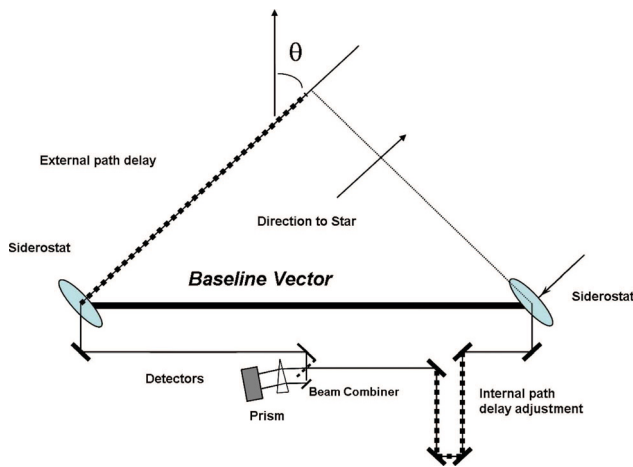


Fig. 2. (Color online) Schematic of a stellar interferometer.

in [5] to obtain the optical path difference (OPD) from the observation data.

Figure 2 shows a schematic of a stellar interferometer. Star light enters the interferometer via the two siderostats, which articulate to steer the light from the star into the interferometer. An imbalance in the internal delay, adjustable by moving a retroreflector in the optical path of one arm, compensates for the imbalance in the external delay; the total delay, then, is balanced. After traveling through the two arms of the interferometer the star light recombines at the beam combiner. A prism spectrally disperses the interference fringes formed by this recombination onto an array of detectors, typically a charge-coupled device (CCD). The angular distance between two stars is inferred from the change in internal delay needed to rebalance the total delay when switching from one star to the other.

Let $I(k)$ denote the dc intensity as a function of wavenumber k , let $V(k)$ denote the frequency dependent visibility, and let $\psi(k)$ be the instrument dependent differential dispersion function. The differential dispersion function arises because of glass mismatches between the two arms of the interferometer, coatings on beam splitters/combiners, etc. Over a single channel the general interference model that is adopted for the observed intensity $Y(x)$ as a function of the vacuum path length difference x in the two arms of the interferometer is

$$Y(x) = \int_{k_-}^{k_+} I(k) \{1 + V(k) \cos(kx + \psi(k))\} dk, \quad (1)$$

where $[k_-, k_+]$ is the passband of the channel. In this model, I and V are generally functions of both the source and the instrument while ψ may be treated as only a function of the instrument. This latter statement requires that either the source is unpolarized or the optical design of the interferometer renders it insensitive to polarization states of the source. The current design of the SIM instrument achieves this

objective by using polarizing beam splitters and an optical design that minimizes the mixing between the two polarization states. The polarizing beams splitters separate the detection of fringes of different polarization states and thus effectively “split” the interferometer into two interferometers measuring fringes generated by the photons corresponding to the two polarization states. Furthermore, to have ψ independent of the source, the source should have negligible phases in its complex visibilities by the van Cittert–Zernike theorem [13] (e.g., either be unresolved by the interferometer or symmetric).

In the companion paper [5], a quasi-monochromatic equivalent to Eq. (1) is produced, which will be the departure point for our investigation. We first need a few definitions and notations. Define

$$\bar{I} = \int_{k_-}^{k_+} I(k) dk, \quad \bar{V} = \frac{1}{\bar{I}} \int_{k_-}^{k_+} I(k) V(k) dk, \quad (2)$$

$$p(k) = \frac{I(k) V(k)}{\bar{I} \bar{V}},$$

and note that $p(k)$ is a density function. The central moments of $p(k)$ are defined as

$$\mu_j = \int_{k_-}^{k_+} (k - \bar{k})^j p(k) dk, \quad j = 0, 1, 2, \dots, \quad (3)$$

where \bar{k} denotes the weighted mean wavenumber:

$$\bar{k} = \int_{k_-}^{k_+} k p(k) dk. \quad (4)$$

The differential dispersion function ψ is written as a sum of an affine part and its nonlinear residual:

$$\psi(k) = a_0 + a_1 k + r(k). \quad (5)$$

This enables a quasi-monochromatic model equivalent to Eq. (1), which can be further reduced to simple parameterizations based on restricting the magnitude of the residual dispersion and the group delay. The two particular reduced models that are studied in this paper are the “second moment” and “fourth moment” models, which are generally valid for describing the interference over group delays on the order of $1 \mu\text{m}$ (as used in modulation methods to identify the phase of the signal). The fourth moment model has been shown to be sufficiently accurate for picometer class OPD estimation when the full passband is divided into four spectral channels, while eight channels are necessary for the second moment model [5]. The second moment model is defined as

$$Y(x) = \bar{I} \left\{ 1 + \bar{V} \left[1 - \frac{\mu_2}{2} u^2 \right] \cos(\bar{k}u + a_0) \right\}, \quad (6)$$

and the fourth moment model is

$$Y(x) = \bar{I} \left\{ 1 + \bar{V} \left[1 - \frac{\mu_2}{2} u^2 + \frac{\mu_4}{24} u^4 \right] \times \cos \left(\bar{k}u + a_0 - \frac{1}{2} \delta u^2 - \frac{\mu_3}{6} u^3 \right) \right\}, \quad (7)$$

where $u = x + a_1$ is the group delay and δ is defined as

$$\delta = \int_{k_-}^{k_+} (k - \bar{k})^2 r(k) p(k) dk. \quad (8)$$

The validity of these approximations are discussed in the companion paper [5].

3. Calibration

The basic strategy for calibrating the parameters required in the fringe estimation models above is to first characterize the system at the “pixel” level. Specifically the mean wavenumber, dispersion term, throughput, and visibility are calibrated for each pixel using a calibration source. In some instances the calibration source and the observed object are the same, and in other scenarios they are different. The models in Eqs. (6) and (7) are then assembled at the spectral channel level (which may bin as many as 20 pixels or as few as 5). In the case that the calibration and observed object are the same, the calibration derived parameters may be used directly. When the two objects are not the same, the calibration parameters must be combined with *a priori* spectral information of the target to produce the models. In these cases a blackbody characterization of the target will be shown to be sufficient.

The SIM spectrometer consists of 80 spectral channels with images illuminating a 1×5 array of pixels for each of the spectral channels. The charge from the five pixels is accumulated for the readout. In an abuse of language, from now on, we will exclusively refer to the readout of this single channel as a pixel. The actual pixel on CCD will be referred as a “physical pixel.” At the pixel level we are also concerned with charge diffusion, diffraction effects, QE of the detector, etc. Thus although adjacent pixels are physically separated, they overlap in terms of their spectral response. Following Lawson [14] the adopted j th-pixel model response is

$$Y_j(x) = \int_0^\infty I_j(k) \{ 1 + V_j(k) \cos(kx + \psi(k)) \} dk, \quad (9)$$

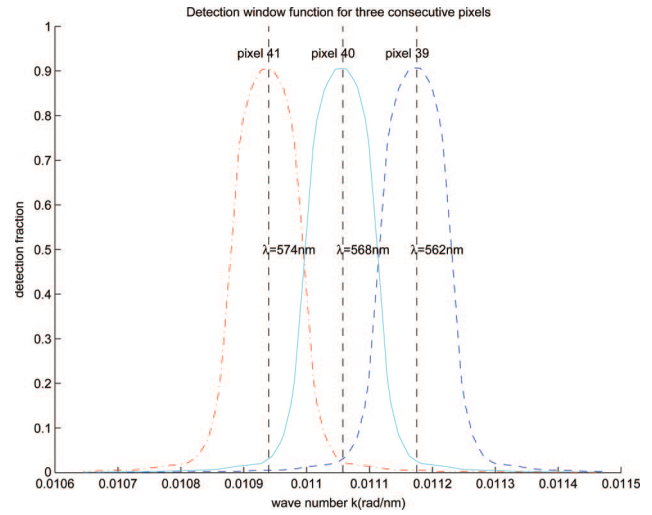


Fig. 3. (Color online) Detection windows for pixels 39, 40, and 41 for the nominal 80 pixel SIM CCD over passband 400–1000 nm.

where I_j is the intensity, a rapidly decreasing function about its peak that includes diffraction and charge diffusion effects. Nominally, the SIM requires that 90% of the energy at a given wavelength be captured in a 3×3 array of physical pixels. In the case of no charge diffusion and infinitesimal diffraction spots, I_j would correspond to a strictly rectangular bandpass that has a cutoff at the pixel boundary. For SIM optics an example of adjacent fringe responses is shown in Fig. 3 using the white-light modeling tool SLIM, which is discussed in more detail in Section 4.

A. Calibration at the Pixel Level

We assume that Eq. (9) may be approximated either by a purely monochromatic or at most a second moment model to correctly capture the response as the fringe of a single pixel is modulated. The correctness of these models depends on the stroke length used in the estimation of the parameters. We will begin with the calibration of a monochromatic model. In Subsection 3.C, a sensitivity analysis of the calibration will be conducted. Subsection 3.D will present a strategy for removing errors due to violating the monochromatic assumption when a long stroke modulation is used.

The basic monochromatic interferometric intensity model is

$$y = A(\bar{k})X + \eta. \quad (10)$$

Here y is the N vector of observed intensities, η is a noise vector (photon statistics, read noise), X is the state vector, and A is the matrix relating the state to the observed intensities. The vector X has components,

$$X = [\bar{I}, \bar{I}\bar{V} \cos(\bar{k}d + \bar{\psi}), \bar{I}\bar{V} \sin(\bar{k}d + \bar{\psi})]^T, \quad (11)$$

where \bar{I} is the dc intensity, \bar{V} is the visibility, \bar{k} is the wavenumber, d is the delay, $\bar{\psi}$ is a dispersion term, and A is an $N \times 3$ matrix:

$$A(k) = \begin{bmatrix} 1 & \cos(ku_1) & -\sin(ku_1) \\ \vdots & \vdots & \vdots \\ 1 & \cos(ku_N) & -\sin(ku_N) \end{bmatrix}. \quad (12)$$

In the system here y is observed and the u_i 's are the introduced delay modulations that are assumed to be accurately measured. This model holds for phase stepping or triangular/sawtooth modulation waveforms. All of the other variables in the system are unknown. The initial objective of the calibration is to determine \bar{k} from the observation vector y . The nonlinear least-squares problem associated with determining \bar{k} is

$$\min_{k, X} |y - A(k)X|^2. \quad (13)$$

Once \bar{k} and X are determined, the variables \bar{I} , \bar{V} , and $\bar{k}d + \bar{\psi}$ can be obtained from X via the inverse of the transformation in Eq. (11). We emphasize that all of these variables are in fact the weighted variables as defined in Eqs. (2)–(4) relative to the system that includes the source and instrument contributions. Observe that this problem is nonlinear in the scalar variable k and is *linear* in the vector X . Least-squares problems of this type can be reduced to a scalar optimization problem in the variable k . Heuristically the argument goes like this: If $k_0 = \bar{k}$ is optimal and $A(k_0)$ has full rank, then the optimal X solves the linear least-squares problem here with $k = k_0$ and is thus determined as

$$X_0 = A(k_0)^\dagger y, \quad (14)$$

where matrix $A(k_0)^\dagger$ is the pseudoinverse [15] of $A(k_0)$. Hence the full nonlinear least-squares problem is equivalent to first solving the *scalar* problem,

$$\min_k |y - A(k)A(k)^\dagger y|^2, \quad (15)$$

and then invoking the linear solution. This heuristic argument is easily made rigorous [16]. For our purposes a great deal has been accomplished since the nonlinear scalar problem is very tractable.

An efficient implementation for solving Eq. (15) is now presented. First recognize that $A(k)A(k)^\dagger$ is the orthogonal projection onto the range of $A(k)$. Denoting this projection as $P(k)$, we see that $|y - P(k)y|^2 = |y|^2 - |P(k)y|^2$. Thus it is sufficient to maximize $|P(k)y|^2$. The computation of the projection $P(k)$ is straightforward and can be accomplished by using the Gram–Schmidt orthogonalization of the columns

of $A(k)$ or by standard Householder transformations to obtain the QR factorization of A , viz. $QR = A$. The orthogonal matrix Q is an $N \times 3$ matrix, and the projection is obtained as $P(k) = QQ^T$. Thus $|P(k)y|^2 = |Q^T y|^2$, and the optimization problem that needs to be solved is simply

$$\max_k |Q^T(k)y|^2. \quad (16)$$

$Q(k)$ will be a very smooth (analytic) function so long as A has constant rank, and consequently should be very well behaved and amenable to any number of one-variable optimization routines [17]. Note that no knowledge of X is required.

In Fig. 4, a normalized version of Eq. (16) is shown for stroke lengths of 2 and 10 μm . In both cases the function is convex, and hence univariate, in a rather large neighborhood of the solution. Observe that the 10 μm stroke provides a functional with much greater curvature than the 2 μm stroke. It can be shown that the curvature of the cost functional increases quadratically with the length of stroke. The detailed proof is provided in Appendix B.

Once \bar{k}_j is obtained within each pixel, in the standard way the other variables, \bar{I}_j , \bar{V}_j , $\bar{k}_j d + \bar{\psi}_j$ are also readily derived. All of these variables are necessary for building the binned channel models in Eqs. (6) and (7). We will first show how the phase $\bar{\phi}_j \equiv \bar{k}_j d + \bar{\psi}_j$ that is computed in each channel is related to the dispersion function $\psi(k)$.

To derive the relationship between $\bar{\phi}_j$ and $\psi(k)$, first observe

$$\bar{\phi}_j = \bar{k}_j u_j + a_0, \quad (17)$$

where a_0 and u_j are related to the differential dispersion function across the j th pixel as $\psi(k) = a_0 + a_1 k + r_0(k)$ with $u_j = d + a_1$ and d denoting the vacuum path length delay. In the companion paper [5] the affine parameters a_0 , a_1 were defined as

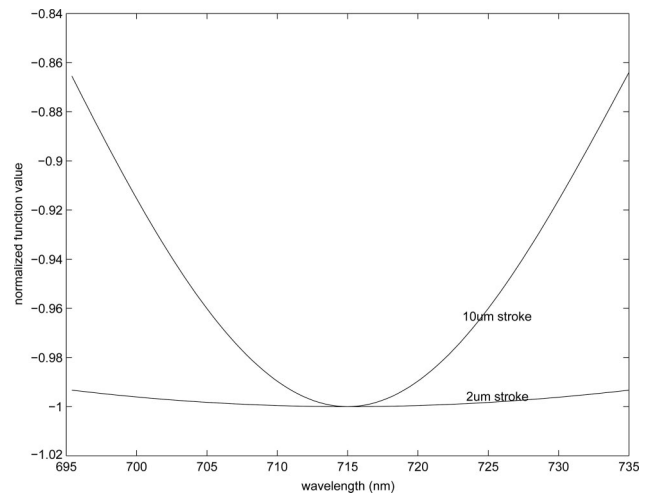


Fig. 4. Cost functionals for stroke lengths of 2 and 10 μm .

$$a_1 = \frac{1}{\mu_2} \int (k - \bar{k}_j) \psi(k) p_j(k) dk,$$

$$a_0 = \int \psi(k) p_j(k) dk - a_1 \bar{k}_j, \quad (18)$$

where p_j is the density function of the j th pixel and μ_2 is its second central moment. Hence,

$$\bar{\phi}_j = \bar{k}_j d + \bar{\psi}_j; \quad \bar{\psi}_j = \int \psi(k) p_j(k) dk. \quad (19)$$

We remark that the a_0, a_1 are also the coefficients of the optimal affine fit to $\psi(k)$ in a weighted least-squares sense [5].

Given the relationship between the $\bar{\phi}_j$'s and the delay d in Eq. (19), the delay may be computed by a linear least-squares fit to the unwrapped phases $\bar{\phi}_j$. An alternative to this is to work directly with the phasors $\{\cos(\bar{\phi}_j), \sin(\bar{\phi}_j)\}$ and computed wavenumbers \bar{k}_j to estimate d [14]. Once the vacuum path length delay d is obtained the weighted average (with respect to the spectrum of the source) of the dispersion function of each pixel may be determined. However, the dispersion function is observable only to within a linear term in wavenumber as a linear dispersion term is indistinguishable from a geometric delay difference.

For smooth differential dispersion functions, over a pixel $\bar{\psi}_j \approx \psi(\bar{k}_j)$. To characterize how close this approximation is consider the Taylor expansion of ψ with remainder: $\psi(k) = \psi(\bar{k}) + \psi'(\bar{k})(k - \bar{k}) + r(k)$ where $r(k) = \psi''(k')(k - \bar{k})^2/2$ for some k' between \bar{k} and k . Then

$$\int \psi(k) p(k) dk = \psi(\bar{k}) + \frac{1}{2} \int \psi''(k')(k - \bar{k})^2 p(k) dk, \quad (20)$$

so that

$$\left| \int \psi(k) p(k) dk - \psi(\bar{k}) \right| \leq \frac{1}{2} \int |\psi''(k')(k - \bar{k})^2 p(k)| dk$$

$$\leq \frac{1}{2} |\psi''|_{\infty} \mu_2. \quad (21)$$

The error above is the spectral dependent error. For the canonical 5 μm BK7 glass imbalance example described in the companion paper [5], $1.5 \times 10^4 \text{ nm}^2 < |\psi''(k)| < 5.5 \times 10^4 \text{ nm}^2$ for wavenumbers corresponding to wavelengths between 400–1000 nm. As the spectrometer disperses the light into 80 channels of roughly equal width in wavenumber, approximat-

ing the density function as rectangular across each pixel's passband, the error in Eq. (21) is on the order of 10^{-5} rad.

B. Assembling the Model at the Channel Level

The parameters of interest in the channel model include the mean wavenumber, second moment, and phase term for the second moment model, and in addition the fourth moment and quadratic phase term for the fourth moment model. There is also a distinction between how the second and fourth moment models are calibrated. For the latter the calibration is done directly on the object, i.e., the star that the fringe estimation algorithm is applied to is also the star used for calibration. In the former case, a calibration star is used and the parameters obtained from it are used in conjunction with *a priori* information of the star that is observed. The fourth moment algorithm is applied to the SIM's guide interferometers, which observe relatively bright stars of approximately seventh magnitude at a high camera frame rate because of certain real time control responsibilities [18]. Because of this high camera frame rate, the spectrometer is divided into just four channels by binning the 80 pixel channels to reduce read noise. The second moment algorithm is applied at a 50 times slower fringe estimation rate using 16 spectral channels.

The first order of business is to compute \bar{k} using the calibrated \bar{I}_j and \bar{V}_j over each pixel:

$$\bar{k} = \frac{\sum_j \int I_j(k) V_j(k) k dk}{\sum_j \int I_j(k) V_j(k) dk} = \frac{\sum_j \bar{k}_j \int I_j(k) V_j(k) dk}{\sum_j \int I_j(k) V_j(k) dk}$$

$$= \frac{\sum_j \bar{k}_j \bar{I}_j \bar{V}_j}{\sum_j \bar{I}_j \bar{V}_j}, \quad (22)$$

where the summation is over the component pixels within the spectral channel. When the calibration and target star are the same, the mean wavenumber for the spectral channel computed here is correct. When they are different, errors arise in two ways. First the pixel wavenumber \bar{k}_j is off because it is spectral dependent and the "wrong" spectrum has been used in its calibration. Second the weights $\bar{I}_j \bar{V}_j$ are also obtained from the "wrong" source. The approximation of the second moment is calculated next. First observe that the true value over the channel is

$$\mu_2 = \frac{\sum_j \int (k - \bar{k})^2 I_j(k) V_j(k) dk}{\sum_j \int I_j(k) V_j(k) dk}. \quad (23)$$

Now,

$$\begin{aligned} \int (k - \bar{k})^2 I_j(k) V_j(k) dk &= \int k^2 I_j(k) V_j(k) dk \\ &\quad - 2\bar{k} \int I_j(k) V_j(k) dk \\ &\quad + \bar{k}^2 \int I_j(k) V_j(k) dk, \end{aligned} \quad (24)$$

$$\begin{aligned} \int k^2 I_j(k) V_j(k) dk &= \int (k - \bar{k}_j)^2 I_j(k) V_j(k) dk \\ &\quad + 2 \int k \bar{k}_j I_j(k) V_j(k) dk \\ &\quad - \bar{k}_j^2 \int I_j(k) V_j(k) dk \\ &= \mu_2^j \bar{I}_j \bar{V}_j + \bar{k}_j^2 \bar{I}_j \bar{V}_j, \end{aligned} \quad (25)$$

where μ_2^j is the second moment over the j th pixel as defined by Eq. (3). Therefore,

$$\mu_2 = \frac{\sum_j (\bar{k} - \bar{k}_j)^2 \bar{I}_j \bar{V}_j}{\sum_j \bar{I}_j \bar{V}_j} + \frac{\sum_j \mu_2^j \bar{I}_j \bar{V}_j}{\sum_j \bar{I}_j \bar{V}_j}. \quad (26)$$

Expression (26) enables us to compute μ_2 for a channel directly from the pixelwise data. The first term gives the dominant contribution to μ_2 while the second term may be small enough to ignore, or as a correction that can be evaluated using approximate values of the pixelwise second moments (e.g., they can be evaluated by assuming a top hat pixel response). Similar considerations lead to the approximations of the third and fourth moments:

$$\mu_3 = \frac{\sum_j (\bar{k} - \bar{k}_j)^3 \bar{I}_j \bar{V}_j}{\sum_j \bar{I}_j \bar{V}_j} + O(|\mu_3^j| + \max_j \mu_2^j |\bar{k} - \bar{k}_j|), \quad (27)$$

$$\begin{aligned} \mu_4 &= \frac{\sum_j (\bar{k} - \bar{k}_j)^4 \bar{I}_j \bar{V}_j}{\sum_j \bar{I}_j \bar{V}_j} + O(|\mu_4^j| + \max_j |\mu_3^j| |\bar{k} - \bar{k}_j| \\ &\quad + \max_j \mu_2^j |\bar{k} - \bar{k}_j|^2), \end{aligned} \quad (28)$$

where μ_3^j and μ_4^j are the third and fourth moments over the j th pixel. Since the passband of a pixel is narrow, μ_3^j and μ_4^j are usually negligible.

We also need to construct approximations of a_0 , a_1 , and δ from the pixelwise data. The computation of δ

depends on the choice for a_0 , a_1 , which we have emphasized before as not being unique. For example we can use a_0 and a_1 from Eq. (18) to have

$$\begin{aligned} a_0 &= \int \psi(k) p(k) dk - \bar{k} a_1 = \frac{\sum_j \int \psi(k) I_j(k) V_j(k) dk}{\sum_j \int I_j(k) V_j(k) dk} - \bar{k} a_1 \\ &= \frac{\sum_j \bar{I}_j \bar{V}_j \bar{\psi}_j}{\sum_j \bar{I}_j \bar{V}_j} - \bar{k} a_1, \end{aligned} \quad (29)$$

$$a_1 = \frac{1}{\mu_2} \frac{\sum_j \bar{I}_j \bar{V}_j (\bar{k} - \bar{k}_j) \bar{\psi}_j}{\sum_j \bar{I}_j \bar{V}_j}, \quad (30)$$

where μ_2 is the estimate from Eq. (26). Thus the pixelwise calibration leads to the precise estimate of a_0 and an approximate estimate of a_1 . Now $r(k) = \psi(k) - a_0 - a_1 k$, and the approximation of δ is

$$\begin{aligned} \delta &= \int (k - \bar{k})^2 r(k) p(k) dk \\ &\approx \frac{\sum_j \bar{I}_j \bar{V}_j (\bar{\psi}_j - a_0 - a_1 \bar{k}_j) (\bar{k}_j - \bar{k})^2}{\sum_i \bar{I}_j \bar{V}_j}. \end{aligned} \quad (31)$$

When the target star and calibration star differ the scenario corresponds to using the second moment model, and the important calibration quantities are \bar{k} and μ_2 . Reasonable estimates for these are obtained using either Eq. (22) and Eq. (26) directly or to retain the calibrated wavenumbers \bar{k}_j from the calibration source and substitute a model of the spectral energy distribution from the target for the $\bar{I}_j \bar{V}_j$ values. The fringe estimation algorithm developed in the companion paper [5] is inherently robust with respect to these two parameters. Examples of the calibration and estimation robustness will be given in Section 4. In Table 1 we compare wavenumbers (wavelengths) and the second moments across eight channels for several stellar types obtained from Pickles [19]. We show the range, mean, and standard deviation of errors between the stellar types across the spectrometer. The channel effective wavenumbers do not vary more than 0.5% for the spectral types shown in the table. The variations of the second moments from star to star are less than a few percent in general.

C. Sensitivity Analysis

Thus far we have ignored the presence of noise in the measurements and potential errors due to deviations from the monochromatic model for the pixelwise fringe. As alluded to in Subsection 3.A there is an advantage to using a long stroke to calibrate the

Table 1. Channel Effective Wavelengths and Second Moments for Different Star Spectra

λ (nm)	Channel 1	Channel 2	Channel 3	Channel 4	Channel 5	Channel 6	Channel 7	Channel 8
A0V	420.60	453.35	493.90	543.18	604.14	681.08	777.67	895.33
F0V	420.69	453.66	494.10	543.88	604.79	681.66	778.92	895.35
G0V	420.61	453.92	494.21	544.35	605.30	682.36	779.54	896.05
K0V	420.69	454.28	494.23	544.97	605.81	682.53	779.89	896.71
M0V	421.68	455.81	493.78	547.05	606.38	683.76	782.08	898.36
Mean	420.85	454.20	494.04	544.68	605.28	682.28	779.62	896.36
Standard deviation	0.46	0.96	0.20	1.47	0.87	1.01	1.61	1.25
Range	1.08	2.46	0.45	3.87	2.24	2.69	4.41	3.03
μ_2 (rad/nm) ²	Channel 1	Channel 2	Channel 3	Channel 4	Channel 5	Channel 6	Channel 7	Channel 8
A0V	1.01×10^{-7}	1.09×10^{-7}	1.22×10^{-7}	1.20×10^{-7}	1.21×10^{-7}	1.21×10^{-7}	1.19×10^{-7}	9.68×10^{-8}
F0V	1.06×10^{-7}	1.11×10^{-7}	1.21×10^{-7}	1.19×10^{-7}	1.21×10^{-7}	1.20×10^{-7}	1.18×10^{-7}	9.80×10^{-8}
G0V	1.15×10^{-7}	1.12×10^{-7}	1.20×10^{-7}	1.18×10^{-7}	1.21×10^{-7}	1.20×10^{-7}	1.19×10^{-7}	9.66×10^{-8}
K0V	1.24×10^{-7}	1.12×10^{-7}	1.20×10^{-7}	1.18×10^{-7}	1.21×10^{-7}	1.20×10^{-7}	1.18×10^{-7}	9.68×10^{-8}
M0V	2.11×10^{-7}	1.14×10^{-7}	1.24×10^{-7}	1.13×10^{-7}	1.23×10^{-7}	1.22×10^{-7}	1.18×10^{-7}	9.78×10^{-8}
Mean	1.31×10^{-7}	1.12×10^{-7}	1.22×10^{-7}	1.18×10^{-7}	1.22×10^{-7}	1.21×10^{-7}	1.18×10^{-7}	9.72×10^{-8}
Standard deviation	4.53×10^{-8}	1.90×10^{-9}	1.53×10^{-9}	2.80×10^{-9}	1.03×10^{-9}	8.32×10^{-10}	3.86×10^{-10}	6.31×10^{-10}
Range	1.10×10^{-7}	5.08×10^{-9}	3.92×10^{-9}	6.95×10^{-9}	2.49×10^{-9}	2.04×10^{-9}	9.64×10^{-10}	1.33×10^{-9}

wavenumbers. However there is simultaneously the disadvantage of straying too far from the monochromatic model assumptions when doing this. Both of these errors will be characterized in this section, and a mitigation strategy will be introduced for calibrating with a long stroke, which leads to a violation of the monochromatic model assumptions. We will first investigate the sensitivity of the monochromatic model calibration scheme to shot and read noise.

From the monochromatic model in Eq. (10) the variations of the parameters and the corresponding signal are related by

$$\delta y_i = \delta I + \delta(IV) \cos(\bar{k}u_i + \bar{\phi}) - \bar{I}\bar{V}\delta k \sin(\bar{k}u_i + \bar{\phi})u_i - \bar{I}\bar{V} \sin(\bar{k}u_i + \bar{\phi})\delta\phi. \quad (32)$$

Defining

$$D \equiv \begin{bmatrix} 1 & \cos(\bar{k}u_1 + \bar{\phi}) & -\sin(\bar{k}u_1 + \bar{\phi}) & -\sin(\bar{k}u_1 + \bar{\phi})u_1 \\ 1 & \cos(\bar{k}u_2 + \bar{\phi}) & -\sin(\bar{k}u_2 + \bar{\phi}) & -\sin(\bar{k}u_2 + \bar{\phi})u_2 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & \cos(\bar{k}u_N + \bar{\phi}) & -\sin(\bar{k}u_N + \bar{\phi}) & -\sin(\bar{k}u_N + \bar{\phi})u_N \end{bmatrix}, \quad (33)$$

$$\delta X \equiv [\delta I, \delta(IV), \bar{I}\bar{V}\delta\phi, \bar{I}\bar{V}\delta k]^T,$$

the relation in vector form is simply

$$\delta y = D\delta X. \quad (34)$$

Note that Eq. (34) may be viewed as a linearized model in the neighborhood of the true value of the intensity, visibility, wavenumber, and phase.

In Appendix C, we find that the column vectors of D are approximately orthogonal to each other for long strokes. It is convenient to rescale the column vectors of D by writing

$$D = V\Lambda, \quad (35)$$

with Λ being a diagonal matrix containing the normalization factors,

$$\Lambda = \text{diag}(\sqrt{N}, \sqrt{N/2}, -\sqrt{N/2}, -\sqrt{N^3\Delta u^2/24}), \quad (36)$$

so that V is approximately an orthogonal matrix. The pseudoinverse is then given by

$$D^\dagger = (D^T D)^{-1} D^T \approx \Lambda^{-1} V^T. \quad (37)$$

See Eqs. (C3), (C4), and (C7) in Appendix C for details.

Now let δy^{err} represent the measurement noise in the observation vector $y = (y_1, \dots, y_N)$ of the signal. Within the linear approximation, its effect on the calibration of the parameters I, V, k, ϕ may be found from a least-squares fit of Eq. (34) to δy^{err} with solution:

$$\delta X = D^\dagger \delta y^{\text{err}}. \quad (38)$$

From Eq. (38) the covariance matrix of δX is

$$E(\delta X \delta X^T) = D^\dagger \text{Cov}(\delta y^{\text{err}}) (D^\dagger)^T \approx \Lambda^{-1} V^T \text{Cov}(\delta y^{\text{err}}) V \Lambda^{-1}, \quad (39)$$

where $\text{Cov}(\delta y^{\text{err}})$ is the covariance matrix of the measurement error δy^{err} . Thus the variance of the i th component of δX is computed as

$$E(\delta X_i^2) = \Lambda_{ii}^{-2} v_i^T \text{Cov}(\delta y^{\text{err}}) v_i, \quad (40)$$

where v_i is the i th column of V [Eq. (C3)].

Now decompose $\text{Cov}(\delta y^{\text{err}})$ into its read and shot noise components:

$$\text{Cov}(\delta y^{\text{err}}) = \sigma_{\text{read}}^2 I_{N \times N} + \bar{I} \text{diag}(1 + \bar{V} \cos(\bar{k}u_1 + \bar{\phi}), \dots, 1 + \bar{V} \cos(\bar{k}u_N + \bar{\phi})), \quad (41)$$

where $I_{N \times N}$ is the $N \times N$ identity matrix. Since each v_i has unit norm, we find that

$$v_i^T \text{Cov}(\delta y^{\text{err}}) v_i = \sigma_{\text{read}}^2 + \bar{I} + \bar{I} \bar{V} v_i^T \text{diag}(\cos(\bar{k}u_1 + \bar{\phi}), \dots, \cos(\bar{k}u_N + \bar{\phi})) v_i. \quad (42)$$

In Appendix E, we show that, for each $i = 1, \dots, 4$:

$$v_i^T \text{diag}(\cos(\bar{k}u_1 + \bar{\phi}), \dots, \cos(\bar{k}u_N + \bar{\phi})) v_i = O(1/N). \quad (43)$$

Now we can compute the relevant variances in Eq. (40), for $i = 1, 2, 3, 4$, using $L = N\Delta u$ and $I_{\text{tot}} = N\bar{I}$:

$$\begin{aligned} E(\delta k^2) &= \frac{1}{[\bar{I}\bar{V}]^2} \left[\frac{24}{NL^2} \right] \{ \sigma_{\text{read}}^2 + \bar{I} \} \\ &= \frac{24\sigma_{\text{read}}^2}{I_{\text{tot}}^2 \bar{V}^2 L \Delta u} + \frac{24}{I_{\text{tot}} \bar{V}^2 L^2}, \end{aligned} \quad (44)$$

$$\begin{aligned} E(\delta \phi^2) &= \frac{2N\sigma_{\text{read}}^2}{(I_{\text{tot}} \bar{V})^2} + \frac{2}{I_{\text{tot}} \bar{V}^2}, \\ E(\delta (IV)^2) &= \frac{2\sigma_{\text{read}}^2}{N} + \frac{2I_{\text{tot}}}{N^2}, \\ E(\delta I^2) &= \frac{\sigma_{\text{read}}^2}{N} + \frac{I_{\text{tot}}}{N^2}. \end{aligned} \quad (45)$$

The important observations to be made from these expressions are that the variance in the wavenumber decreases quadratically with the stroke length in the photon noise dominated case, as would be expected in the calibration procedure, and the variances of the other parameters are independent of the stroke length. Thus it is desirable to maintain as long a stroke length as possible to estimate the wavenumber. However as will be discussed in Subsection 3.D there are additional considerations.

D. Fringe Model Error Effects and Corrections

So far we have assumed that the fringe signal is developed by interfering monochromatic light. For SIM, the light sources are the stars. The nonzero bandwidth of the light source and detector produces a fringe pattern with an envelope that is inversely proportional to the bandwidth. Ignoring the envelope leads to a model error, even at the pixel level, that in turn causes calibration errors. When a long stroke is used, the effect of the envelope becomes even more important for the calibration of visibility and phase dispersion. (See the sensitivity study in Appendix C.) In this section we will show how to compensate for this effect.

Qualitatively, the second moment model in Eq. (6),

$$Y(x) \approx \bar{I} \left\{ 1 + \bar{V} \left[1 - \frac{1}{2} \mu_2 (x + a_1)^2 \right] \cos(\bar{k}x + \bar{\psi}) \right\} \quad (46)$$

is a monochromatic signal with a quadratic envelope characterized by the group delay $u = x + a_1$ and the second moment μ_2 , which determine the center and the width of the envelope, respectively. When the bandwidth is narrow so that $\mu_2 u^2 \ll 1$, as it is for a pixel, the effect of the envelope factor is small compared with that of the monochromatic factor even for relatively long stroke fringe data. Therefore, we will treat the fringe envelope portion as a perturbation to the monochromatic signal.

Incorporating the envelope model, the counterpart to the monochromatic least-squares problem in Subsection 3.A is

$$\begin{aligned} &\min_{k,X} |y - [A(k)X + \delta y^{\text{env}}]|^2, \\ \delta y_i^{\text{env}} &= -\frac{1}{2} \bar{I} \bar{V} \mu_2 (u_i + d + a_1)^2 \cos(\bar{k}u_i + \bar{k}d + \bar{\psi}), \\ &i = 1, 2, \dots, N, \end{aligned} \quad (47)$$

where y is the fringe measurement, $A(k)X$ is the monochromatic signal, and δy^{env} is the perturbation due to the envelope effect. The main idea is to introduce a projection operator P that annihilates the perturbation, i.e., $P\delta y^{\text{env}} = 0$. Because P projects vectors onto the subspace orthogonal to δy^{env} , k and x may be obtained from the alternate problem,

$$\min_{k,X} |Py - PA(k)X|^2, \quad (48)$$

in which the envelope effect has been removed.

There are two issues to explore with the introduction of projection P . First, because we do not know all of the parameters exactly, it is not possible to completely annihilate δy^{env} , hence there is always a residual error. Also, P will in general annihilate some of the signal $A(k)X$, so there will be some reduction in the SNR as well. These details are fleshed out below.

Observe from Eq. (47) that the envelope signal δy^{env} is spanned by $e_0 = \cos(\bar{k}u_i + \bar{\phi})$ and the two vectors e_1, e_2 :

$$\begin{aligned} e_1 &\equiv u_i \cos(\bar{k}u_i + \bar{\phi}), & e_2 &\equiv u_i^2 \cos(\bar{k}u_i + \bar{\phi}), \\ &i = 1, 2, \dots, N. \end{aligned} \quad (49)$$

Since the vector e_0 is indistinguishable from the monochromatic signal, the projection operator is developed to annihilate only e_1 and e_2 , i.e., $Pe_1 = Pe_2 = 0$. The presence of e_0 causes a small error in the visibility estimate. Constructing the vectors e_1 and e_2 requires estimates of \bar{k} and $\bar{\phi}$. Good estimates of these parameters can be obtained from the monochromatic

model by ignoring the fringe envelope contribution. But since there are errors in these estimates the envelope may not be completely annihilated by P . In general, the wavenumber estimate is insensitive to the fringe envelope error [See Eq. (C10) in Appendix C] so that the errors in e_1 and e_2 are mainly attributable to the estimation error $\bar{\phi}$. Thus to ensure an adequate suppression of the envelope signal even when $\bar{\phi}$ is not precisely known, we include a third vector e_3 defined

$$e_3 \equiv u_i^2 \sin(\bar{k}u_i + \bar{\phi}), \quad (50)$$

as another dimension for the error signal space and require P to satisfy the additional condition $Pe_3 = 0$. Although it is tempting, we do not include the vector $\sin(\bar{k}u_i + \bar{\phi})u_i$ in this development for two reasons. First, e_1 is smaller than e_2 for longer strokes where the fringe envelope is more important. Since $\bar{\phi}$ is already known to first order, the error parallel to the vector $\sin(\bar{k}u_i + \bar{\phi})u_i$ is very small. Second, the signal parallel to $\sin(\bar{k}u_i + \bar{\phi})u_i$ should not be projected away because it affects the determination of the wavenumber \bar{k} . The main signal is $\cos(\bar{k}u_i + \bar{\phi})$, varying \bar{k} moves in the signal space along vector $\sin(\bar{k}u_i + \bar{\phi})u_i$. Thus if we work in a subspace orthogonal to $\sin(\bar{k}u_i + \bar{\phi})u_i$, we completely lose the sensitivity to \bar{k} because its variation has no effect on the fitting.

Treating e_1 , e_2 , and e_3 as column vectors, the QR factorization,

$$\mathcal{Q}\mathcal{R} = [e_1 \ e_2 \ e_3], \quad (51)$$

yields the orthogonal matrix \mathcal{Q} whose columns span the same subspace as e_1 , e_2 , e_3 , which is our approximation of the fringe envelope signal subspace. The required projection P onto the orthogonal complement is then

$$P = I - \mathcal{Q}\mathcal{Q}^T. \quad (52)$$

Thus we may replace the matrix $A(k)$ in Eq. (12) with $PA(k)$ to suppress the fringe model error corresponding to the envelope. Because the envelope is a perturbation to the main signal, the initial estimates of \bar{k} , $\bar{\phi}$, \bar{I} , and \bar{V} using a monochromatic algorithm are in the neighborhood of the true effective wavenumber, phase, intensity, and visibility. This enables us to take a perturbative approach as an alternative, i.e., using the linearized model in Eq. (34) to determine the correction to the initial estimates. Mathematically, we solve for δX in the least-squares sense from

$$P\delta y = PD\delta X, \quad (53)$$

where δy is the residual signal after subtracting the monochromatic signal portion corresponding to the initial estimated parameters. The solution is again expressed in terms of a pseudoinverse as

$$\delta X = (PD)^+ P\delta y. \quad (54)$$

A sensitivity analysis, very similar to the one performed in the monochromatic case reveals that the penalty for including the projection operator is a 9/4 increase in variance of the visibility and phase estimates. However, there is no increase in the variance of the estimate of the wavenumber. The details of this analysis are worked out in Appendix C.

4. Results

We applied the calibration and OPD estimation algorithms to data from simulations and also to data from the testbed. The simulations show the expected performance. The stroke length dependency of the wavenumber calibration error is as derived in Subsection 3.C, and the OPD estimation errors in end-to-end simulations are in single digit picometers meeting the SIM requirement on the systematic error considered here. The results also show the efficacy of the projection technique, especially at improving the visibility and phase dispersion calibration accuracy.

Analysis of the testbed data demonstrates the accuracy of the wavelength calibration and a significant reduction of the systematic error in OPD estimation. Our algorithm also yields results meeting the performance metric [4].

A. Simulations

We have developed a modular, general purpose numerical model, the SLIM [11] to simulate both the physics of the interferometer and the processing of data from the interferometer using the estimation algorithms. The two aspects (physics and processing) are simulated by separate blocks of the model. The physics block generates simulated CCD output and simulated metrology readings. The estimation block takes the CCD output and metrology readings and generates delay estimates.

One of the functions of the physics block is to simulate the interference and detection of the starlight in the interferometer. Each point on the CCD receives interfered starlight with an intensity that depends on the stellar spectrum, the spectral visibility of the interferometer, and on the fringe phase at that wavelength. The fringe phase depends on the position of the modulator, on additional geometric path differences due to the orientation and deformation of the spacecraft, and on nongeometric optical path differences, such as those caused by propagation through dispersive optical elements. During an exposure interval of the CCD, the number of photoelectrons generated in a region of finite area (e.g., a pixel) is proportional to the triple integral, over that area and over time, of the instantaneous local intensity. The diffraction effect is incorporated as an extra multiplicative factor in the integrand describing the diffraction pattern. This triple integral is calculated in the physics block using a quadrature algorithm. In simulations in which the local intensity is assumed to be independent of the cross-spectral direction (the direction on the surface of the CCD perpendicular to the

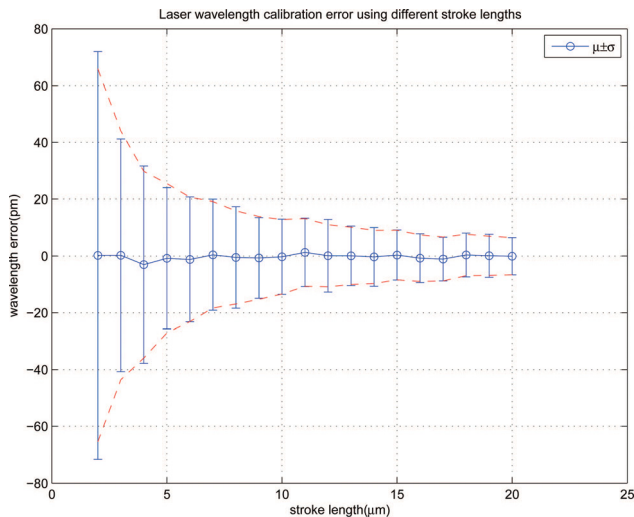


Fig. 5. (Color online) Wavelength calibration error using different stroke lengths; the dashed curves show the standard deviation that is inversely proportional to the stroke length using the longest stroke length standard deviation as the scale.

direction in which the fringe light is spectrally dispersed), or in which the cross-spectral variation in intensity is unimportant, a double integral is evaluated instead. Another function of the physics block is

to simulate the operation of the metrology system. In the SIM, internal metrology measurements are made in each arm at a rate of 320 kHz, the heterodyne frequency, and accumulated in the internal metrology card over (62.5 μ s) intervals corresponding to the 16 kHz metrology read rate. In SLIM the accumulation of metrology measurements during the read interval is modeled using a quadrature over time.

To assess the efficacy of the calibration and OPD estimation algorithms, we applied these algorithms to simulated fringe measurements generated using SLIM. The simulation used star spectra from the Pickles database [19] a uniform visibility of 0.8, and a phase dispersion curve corresponding to a 5 μ m imbalance, between the two arms of the interferometer, in the total thickness of the fused silica transmissive optics.

Figure 5 shows the wavelength calibration error due to shot noise as a function of the stroke length for a laser source with the total number of detected photoelectrons fixed at 10^7 . The error is inversely proportional to the stroke length.

Figure 6 shows the calibration results for the intensity, visibility, wavelength, and phase dispersion for an arbitrarily chosen channel spanning the wavelength range of 845–870 nm. We used the least-squares algorithm both with and without projection.

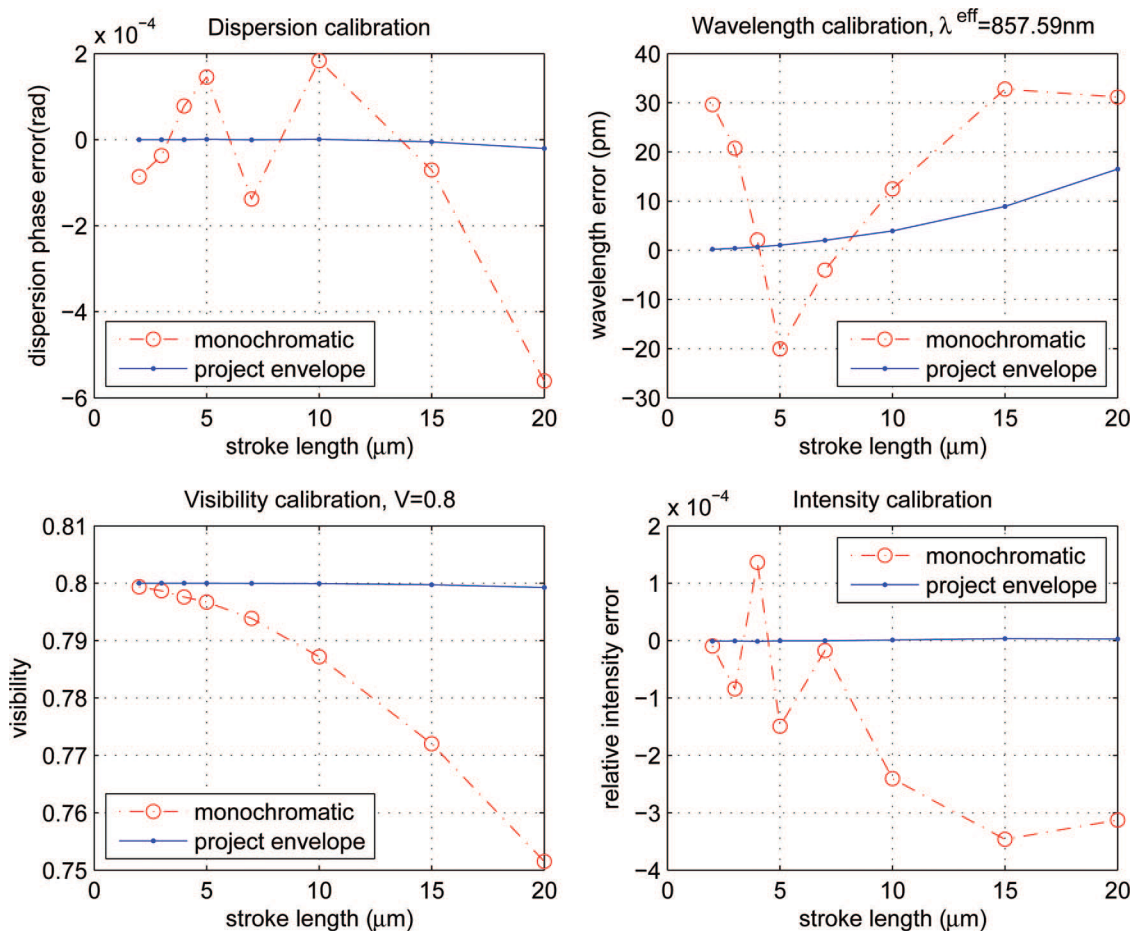


Fig. 6. (Color online) Calibration results with and without the envelope projection technique.

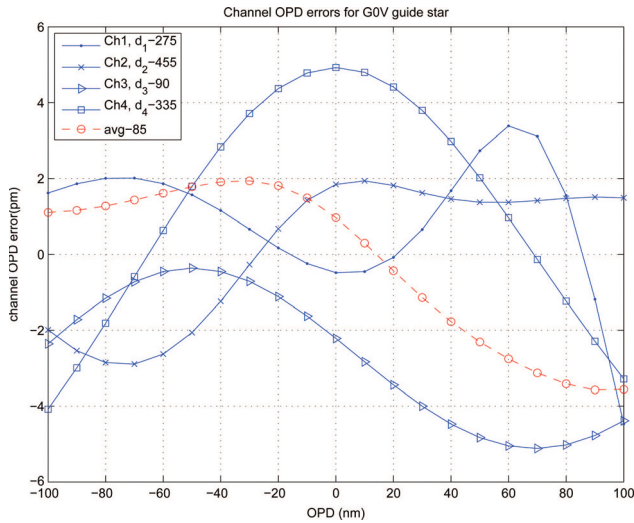


Fig. 7. (Color online) OPD error from an end-to-end simulation for the guide interferometer observation scenario, calibration is performed with noise on.

Without using the the projection technique, the errors shown are consistent with sensitivity study results in Eqs. (C10). The stroke length dependency is much smaller, especially for the visibility and phase dispersion, when we apply the envelope projection technique. The increasing error in the wavelength calibration at long stroke lengths is due to the inadequacy of the quadratic fringe envelope model at large delays.

We present end-to-end run results in Figs. 7 and 8. Figure 7 shows the channel delay estimates over an OPD range of ± 100 nm for the guide interferometer observation scenario. The calibration is done in the presence of noise using a stroke length of $10 \mu\text{m}$ and a seventh magnitude main sequence G0 spectrum for a duration of 60 s. The noise causes hundreds of picometers of phase dispersion calibration error, which in turn causes hundreds of picometers OPD error.

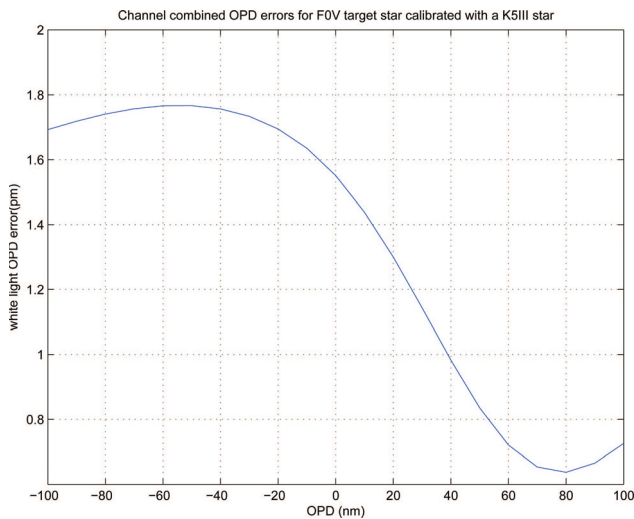


Fig. 8. (Color online) OPD error from an end-to-end simulation for the science observation scenario.

These OPD errors are irrelevant, however, because only changes in the OPD estimate error affect the astrometric performance, and we have subtracted channel-dependent offsets (shown in the legend) from channel delays d_1 through d_4 so that all the curves fit in one plot. The OPD variations shown in Fig. 7 are within 10 pm for ± 100 nm OPD range. We show the combined channel delay for a science interferometer in Fig. 8. For science observations, it is impractical to perform a calibration on each star. Because we use relatively narrow channels (16 spectral channels), the OPD estimation algorithm is not sensitive to errors in our knowledge of the target star spectrum. To demonstrate this insensitivity, we simulated phase estimation performed assuming a normal K5 giant star spectrum on simulated fringes generated using a main sequence F0 star. The total delay error is under 2 pm.

B. Testbed Results

We also analyzed data from the MAM testbed, which was a SIM-like interferometer illuminated by a simulated star light source [3]. The light source produced a mixture of light from a filament and laser light. A CCD was used to detect the combined fringes in 40 spectral channels over a wavelength range of 600–1000 nm. We dithered the OPD with a triangle wave and simultaneously took CCD data at a rate of 500 Hz.

Calibration data were taken using a peak-to-peak dither amplitude of $23 \mu\text{m}$ and a dither period of 10 s. As discussed in Subsection 3.C, using a longer stroke is expected to improve the sensitivity of the wave-number calibration. To test this, we mimicked the effect of shorter dither strokes by discarding data from both ends of each dither stroke. Figure 9 shows the laser wavelength calibration error as a function of the stroke length. The error bars show standard deviations of the calibrated wavelength for 180 strokes

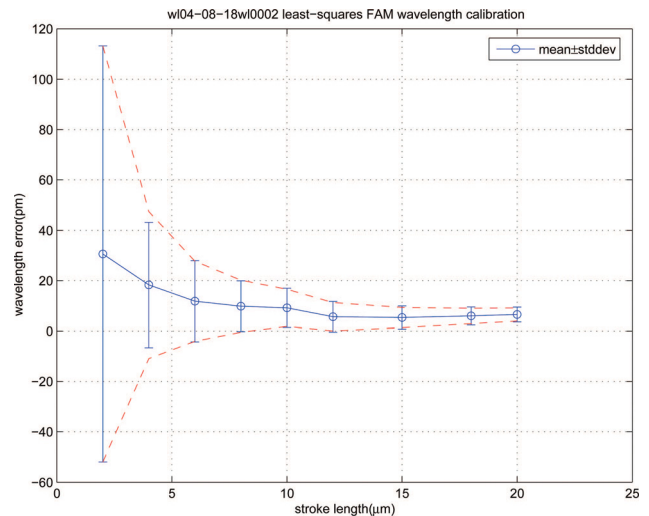


Fig. 9. (Color online) Wavelength calibration error using different stroke lengths, the longest stroke contains $N = 2201$ measurements.

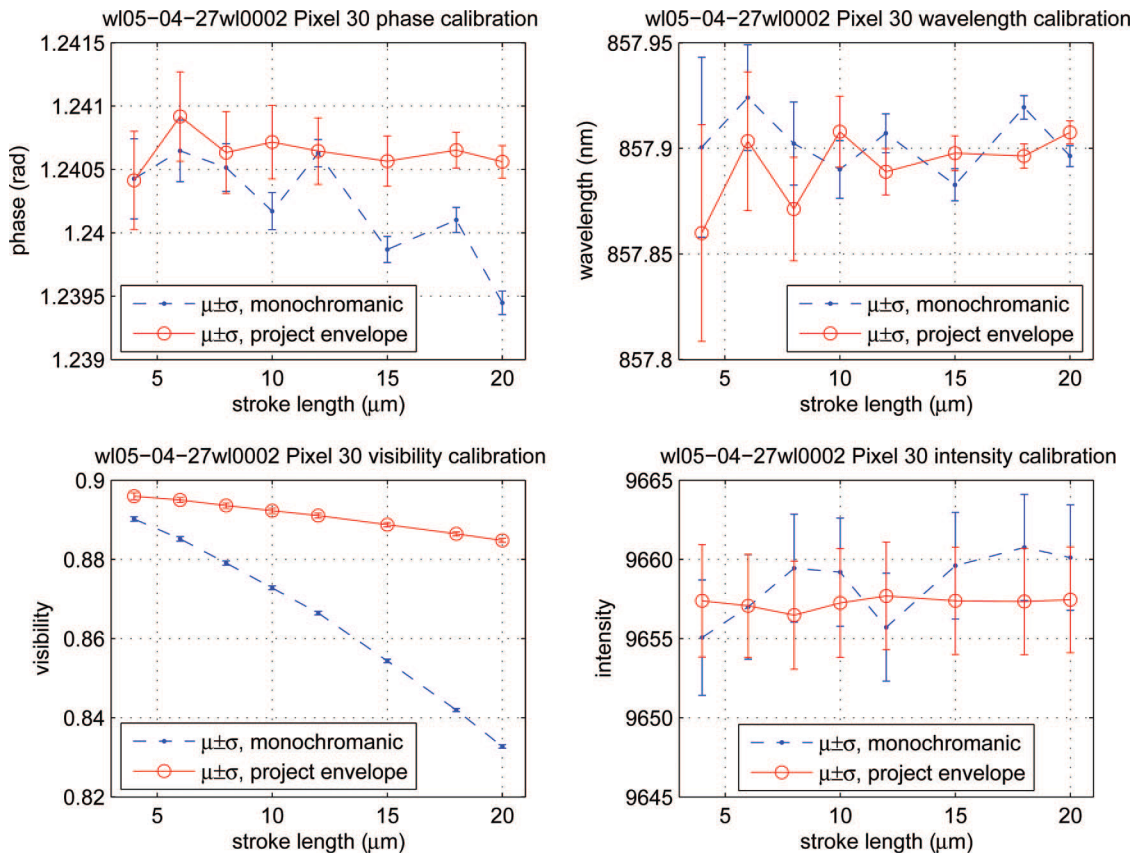


Fig. 10. (Color online) MAM pixel calibration result using and without using the envelope projection technique.

of data; the dashed curves show the standard deviation profile expected from Eq. (44), with $I \propto L$, scaled to the sample standard deviation for a $2 \mu\text{m}$ stroke length. The wavelength error is the deviation from the true laser wavelength, 659.543 nm . The results show that the least-squares algorithm determines the wavelength to an accuracy better than 10 pm . Because the full-aperture metrology (FAM) laser light was generated by doubling the metrology laser, and the metrology laser was used to monitor the OPD during the calibration, this technique is insensitive to laser wavelength drift. It remains sensitive, however, to noise sources such as detection noise. This study shows that the sensitivity to noise decreases when the stroke length increases. The small bias shown in Fig. 9 is probably due to a small amount of light from the MAM white-light source, which is also detected.

Figure 10 shows the calibration results for the intensity, visibility, wavelength, and phase dispersion for fringe signals detected by a MAM pixel (pixel 30) centered at the wavelength 858 nm . We used the least-squares algorithm both with and without projection. We see similar improvement using the envelope projection technique on reducing the stroke length dependency.

Figure 11 shows the OPD estimation errors using the fourth moment, the second moment, and the simple monochromatic algorithms as functions of the actual OPD. The data run used a dither amplitude of $1.3 \mu\text{m}$ and a frequency of 25 Hz . The center of the

dither was changed by 20 nm every 30 s following a staircase profile first decreasing from -20 nm to -80 nm , then increasing to $+80 \text{ nm}$, and then decreasing again to 0 nm . We analyzed fringe data collected in the wavelength range $720\text{--}970 \text{ nm}$. Calibration data were obtained in a previous run over an OPD range of $10 \mu\text{m}$. In MAM, the white-light OPD could drift over many nanometers in a few minutes.

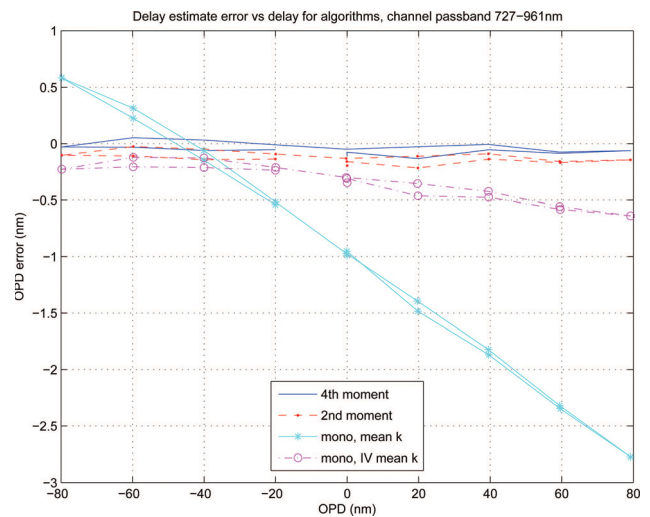


Fig. 11. (Color online) OPD errors using different algorithms for OPD range $\pm 80 \text{ nm}$.

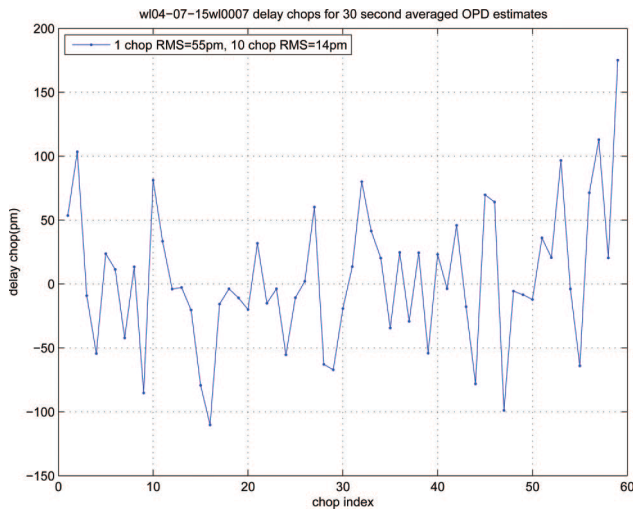


Fig. 12. (Color online) OPD chops for average delay over 30 s.

This drift may be estimated as the mean of the OPD estimates obtained by applying the monochromatic algorithm to the signal from each pixel. Figure 11 shows the broadband OPD estimates minus this estimated drift. The curve labeled “mono, mean k” was obtained using the center of the spectral channel as the effective wavenumber. The error is a few nanometers and is mainly linear in the OPD. The large slope shows that the center wavenumber is a poor choice for the effective wavenumber. The curve labeled “mono, IV mean k” was obtained using instead the “IV” weighted mean wavenumber as defined in Eq. (22). This is a better choice for the effective wavenumber, but it still yields delay estimates with errors of hundreds of picometers. These errors are reduced significantly when using the second or the fourth moment algorithm. The OPD errors for the fourth moment algorithms are in the range of the noise level of 50 pm. The second moment algorithm yields OPD estimates with a small systematic constant error of ~ 100 pm due to ignoring the higher moments in the second moment fringe model. The channel bandwidth here is comparable to that in the SIM guide interferometers. The science interferometer has a narrower bandwidth, and the small systematic error for the second moment algorithm is insignificant.

Figure 12 displays the reference/target (RT) chops as the pairwise differences between the mean OPD errors over consecutive 30 s observations, where the OPD error is the difference between the white-light OPD and the internal metrology [4]. The ten chop deviation calculated here is 14 pm, which is very close to the best performance (approximately 10 pm) achieved using narrowband fringes on the MAM testbed [4].

5. Concluding Remarks

Because the SIM will detect, in as few as four spectrometer channels, starlight in a spectral bandwidth that spans more than one octave, simple algorithms that approximate the detected signal as being pro-

duced by monochromatic light may introduce unacceptable astrometric errors. We have derived phase estimation algorithms in which the product of the visibility and the spectrum of the detected light is represented in terms of moments. The second and fourth moment representations are sufficiently exact, in the science and guide interferometer, respectively, that the remaining estimation error is acceptable.

Both of these algorithms require some knowledge of the star spectrum, of the instrument throughput and phase dispersion, and of the spectral responsivity of each spectrometer channel. We have developed a calibration procedure and algorithm for obtaining this information. The procedure involves capturing the interference fringes while sweeping the OPD through a relatively large change, and the algorithm involves fitting the star spectrum and the instrument characteristics to the resulting data. This fit is both computationally efficient and robust because the model for the interference fringes is nonlinear in only one of the parameters used to fit the data. Both simulations and the analysis of experimental data from the MAM testbed confirm that the calibration procedure and algorithm produce accurate, repeatable data, and that when they are used with the phase estimation algorithms, the residual systematic errors are small.

Because the MAM optical design was sensitive to polarization states of the source, it was not suitable for measuring the error caused by a change in the simulated star spectrum. For this reason, we have presented only simulated results for this error. A new interferometer, the Spectral Calibration Development Unit (SCDU) is currently being built. Because the starlight paths lie nearly in a plane in SCDU and a polarizing beam splitter is used, the polarization-dependent delay is better controlled, and SCDU is expected to provide the opportunity to test the effect of switching between two or more simulated stellar spectra.

Appendix A: Some Useful Sums and Their Asymptotic Behaviors for Long Stroke Length

In this section, we define and compute some useful sums to facilitate the sensitivity analysis for long stroke length in the later appendices. To simplify the analysis, we assume a uniform and symmetric modulation, $u_i = (i - (N + 1)/2)\Delta u$ with Δu being the step size. Define

$$\begin{aligned} C_l(k, \Delta u) &\equiv \sum_{i=1}^N u_i^l \cos(ku_i), \\ S_l(k, \Delta u) &\equiv \sum_{i=1}^N u_i^l \sin(ku_i). \end{aligned} \quad (\text{A1})$$

By symmetry $u_{N+1-i} = -u_i$,

$$C_l = S_{l+1} = 0 \quad \text{for odd } l. \quad (\text{A2})$$

$C_0(k, \Delta u)$ may be computed as

$$C_0(k, \Delta u) = \sum_{i=1}^N \cos(ku_i) = \sum_{i=1}^N \frac{\sin[(i - N/2)k\Delta u] - \sin[(i - N/2 - 1)k\Delta u]}{2 \sin(k\Delta u/2)} = \frac{\sin(Nk\Delta u/2)}{\sin(k\Delta u/2)}. \quad (\text{A3})$$

Using the following relations,

$$\begin{aligned} S_{2l+1}(k, \Delta u) &= -\frac{\partial C_{2l}(k, \Delta u)}{\partial k}, \\ C_{2l+2}(k, \Delta u) &= \frac{\partial S_{2l+1}(k, \Delta u)}{\partial k}, \end{aligned} \quad (\text{A4})$$

we compute

$$\begin{aligned} S_1(k, \Delta u) &= -\frac{\partial C_0}{\partial k} = \Delta u \left[\frac{\sin(Nk\Delta u/2)\cos(k\Delta u/2)}{2 \sin^2(k\Delta u/2)} - \frac{N \cos(Nk\Delta u/2)}{2 \sin(k\Delta u/2)} \right], \\ C_2(k, \Delta u) &= \frac{\partial S_1(k, \Delta u)}{\partial k} = \Delta u^2 \left[\frac{(N^2 + 1)\sin(Nk\Delta u/2)}{4 \sin(k\Delta u/2)} + \frac{N \cos(Nk\Delta u/2)\cos(k\Delta u/2)}{2 \sin^2(k\Delta u/2)} - \frac{\sin(Nk\Delta u/2)}{2 \sin^3(k\Delta u/2)} \right]. \end{aligned} \quad (\text{A5})$$

As far as $k\Delta u$ is not very close to zero, we have the $C_0 \sim 1$, $S_1 \sim N$, $C_2 \sim N^2$ and so on. Taking each derivative gives one more power of N , therefore, the asymptotic behavior for C_l and S_l for large N is

$$\begin{aligned} C_l(k, \Delta u) &\sim N^l \Delta u^l \quad \text{for even } l, \\ S_l(k, \Delta u) &\sim N^l \Delta u^l \quad \text{for odd } l. \end{aligned} \quad (\text{A6})$$

For $k = 0$, using the relations in Eq. (A4), we have

$$C_{2l}(0, \Delta u) = (-1)^l \frac{\partial^{2l}}{\partial k^{2l}} C_{2l}(k, \Delta u) \Big|_{k=0}. \quad (\text{A7})$$

Therefore, up to the factor $(-1)^l$, $C_{2l}(0, \Delta u)$ is the $2l$ th Taylor expansion coefficient of $C_0(k, \Delta u)$ in powers of k . For large N , expanding expression (A3) in powers of k , we find the dominant coefficient for the term k^{2l} is

$$\frac{1}{(2l+1)!} (-1)^l N^{2l+1} \left(\frac{k\Delta u}{2} \right)^{2l}. \quad (\text{A8})$$

Therefore,

$$\begin{aligned} C_{2l}(0, \Delta u) &\approx (2l)! \frac{1}{(2l+1)!} (-1)^l N^{2l+1} \left(\frac{\Delta u}{2} \right)^{2l} \\ &= \frac{N^{2l+1} \Delta u^{2l}}{(2l+1)2^{2l}}. \end{aligned} \quad (\text{A9})$$

In view of definition (A1), taking $k \rightarrow 0$, we get

$$C_{2l}(0, \Delta u) = \sum_{i=1}^N u_i^{2l}, \quad (\text{A10})$$

and thus the large N behavior:

$$\sum_{i=1}^N u_i^{2l} \approx \frac{N^{2l+1} \Delta u^{2l}}{(2l+1)2^{2l}}. \quad (\text{A11})$$

Appendix B: Cost Functional Curvature for Wavelength Calibration

In this appendix, we derive the relation between stroke length and the wavelength calibration sensitivity and prove that the curvature or the second derivative of the cost functional is proportional to the square of the stroke length.

The cost functional expressed in Eq. (13) in Subsection 3.A, may be written as

$$\gamma(k) = |P_{\perp}(k)y|^2 = y^T P_{\perp}(k)^2 y, \quad (\text{B1})$$

where $P_{\perp}(k) = I_{N \times N} - P(k)$ is the orthogonal projection onto the orthogonal complement of the range of $A(k)$. Here $I_{N \times N}$ represent an $N \times N$ identity matrix. We recall that $P(k) = A(k)A(k)^{\dagger}$. Assume y is measured without noise so that $y = A(k_0)X_0$ for the correct wavenumber k_0 with phasor $X_0 = [\bar{I}, \bar{I}V \cos(k_0 d + \bar{\psi}), \bar{I}V \sin(k_0 d + \bar{\psi})]^T$. We will compute $\gamma''(k_0)$ and determine its stroke length dependence.

Taking the first derivative with respect to k gives

$$\begin{aligned} \gamma'(k) &= y^T [P_{\perp}'(k)P_{\perp}(k) + P_{\perp}(k)P_{\perp}'(k)]y \\ &= 2y^T P_{\perp}'(k)P_{\perp}(k)y, \end{aligned} \quad (\text{B2})$$

where we have taken the transpose of the second term and used that $P(k)$ is symmetric. Taking derivatives again, we get

$$\gamma''(k) = 2y^T [P_{\perp}''(k)P_{\perp}(k) + 2P_{\perp}'(k)P_{\perp}'(k)]y. \quad (\text{B3})$$

Since $P_{\perp}(k_0)y = P_{\perp}(k_0)A(k_0)X_0 = 0$,

$$\gamma''(k_0) = 2y^T P_{\perp}'(k_0)P_{\perp}'(k_0)y. \quad (\text{B4})$$

We compute

$$\begin{aligned} P_{\perp}'(k_0)y &= \left[A'(k_0)A(k_0)^{\dagger}A(k_0) + A(k_0)\frac{d}{dk}(A(k)^{\dagger}) \Big|_{k=k_0} A(k_0) \right] X_0 \\ &= (1 - A(k_0)A(k_0)^{\dagger})A'(k_0)X_0 \\ &= (1 - P(k_0))A'(k_0)X_0 \\ &= P_{\perp}(k_0)A'(k_0)X_0, \end{aligned} \quad (\text{B5})$$

where we have used

$$A(k_0)^\dagger A(k_0) = I_{3 \times 3}, \quad (\text{B6})$$

and the following relation derived by taking its derivative:

$$\left. \frac{d}{dk} (A(k)^\dagger) \right|_{k=k_0} A(k_0) + A(k_0)^\dagger A'(k_0) = 0. \quad (\text{B7})$$

Therefore,

$$\gamma''(k_0) = |P_\perp(k_0)z|^2, \quad (\text{B8})$$

where

$$z = A'(k_0)X_0 = -\bar{I}\bar{V}(u_1 \sin(k_0 u_1 + \bar{\phi}), \dots, u_N \sin(k_0 u_N + \bar{\phi})). \quad (\text{B9})$$

The results of Eq. (C5) in Appendix C show that z (proportional to v_4) is nearly orthogonal to the range of $A(k_0)$ (spanned by v_i , $i = 1, 2, 3$). [See Eq. (B11).] Hence $|z|^2 \approx |P(k_0)^\perp z|^2$. With a uniform sampling and setting $\Delta u = (u_{i+1} - u_i)$, we find from Eq. (C5),

$$|z|^2 = O(\bar{I}^2 \bar{V}^2 N^3 \Delta u^2) = O(I_{\text{tot}}^2 \bar{V}^2 N \Delta u^2), \quad (\text{B10})$$

where $I_{\text{tot}} \equiv \bar{I}N$ represents the total number of photoelectrons observed in one stroke. Thus $\gamma''(k_0)$ increases linearly with the number of steps when the step length is fixed, and quadratically with respect to the step length when the number of steps is fixed. Furthermore, since $|y|^2 = O(I_{\text{tot}}^2 \bar{V}^2 / N)$ and the stroke length is $L = N \Delta u$, we find that $\gamma''(k_0) / |y|^2 = O(L^2)$ and thus increases quadratically with length of the stroke. This explains the differences in the curvature between the two functionals in Fig. 4.

Appendix C: Long Stroke Behavior of the Calibration Algorithm Sensitivity

In this section, we provide the details of the sensitivity study for long stroke in Subsection 3.C. We will continue to assume a uniform and symmetric modulation, $u_i = (i - (N + 1)/2)\Delta u$ with Δu being the step size. We will study the asymptotic behaviors of D , $D^T D$, and D^\dagger for long stroke length.

Defining v_i , $i = 1, 2, 3, 4$,

$$\begin{aligned} v_1^i &= \frac{1}{\sqrt{N}}, & v_2^i &= \sqrt{\frac{2}{N}} \cos(\bar{k}u_i + \bar{\phi}), \\ v_3^i &= \sqrt{\frac{2}{N}} \sin(\bar{k}u_i + \bar{\phi}), \\ v_4^i &= \sqrt{\frac{24}{N^3 \Delta u^2}} u_i \sin(\bar{k}u_i + \bar{\phi}) \end{aligned} \quad (\text{C1})$$

enables to write D as

$$D = V\Lambda, \quad (\text{C2})$$

where V and Λ are given by

$$V \equiv [v_1, v_2, v_3, v_4], \quad (\text{C3})$$

$$\Lambda = \begin{bmatrix} \sqrt{N} & 0 & 0 & 0 \\ 0 & \sqrt{\frac{N}{2}} & 0 & 0 \\ 0 & 0 & -\sqrt{\frac{N}{2}} & 0 \\ 0 & 0 & 0 & -\sqrt{\frac{N^3 \Delta u^2}{24}} \end{bmatrix}. \quad (\text{C4})$$

Using the symmetry $u_{N+1-i} = -u_i$ and the asymptotic behaviors in Eq. (A6) of C_l , S_l , and the sums (A11), we evaluate the inner products of the vectors v_i , $i = 1, 2, 3, 4$ as

$$\begin{aligned} \sum_{i=1}^N (v_1^i)^2 &= 1, \\ \sum_{i=1}^N (v_2^i)^2 &= \frac{1}{N} \sum_{i=1}^N [1 + \cos(2\bar{k}u_i + 2\bar{\phi})] \\ &= 1 + \frac{1}{N} \sum_{i=1}^N \cos(2\bar{k}u_i) \cos(2\bar{\phi}) = 1 + O\left(\frac{1}{N}\right), \end{aligned}$$

$$\begin{aligned} \sum_{i=1}^N (v_3^i)^2 &= \frac{1}{N} \sum_{i=1}^N [1 - \cos(2\bar{k}u_i + 2\bar{\phi})] \\ &= 1 - \frac{1}{N} \sum_{i=1}^N \cos(2\bar{k}u_i) \cos(2\bar{\phi}) = 1 + O\left(\frac{1}{N}\right), \end{aligned}$$

$$\begin{aligned} \sum_{i=1}^N (v_4^i)^2 &= \frac{12}{N^3 \Delta u^2} \sum_{i=1}^N u_i^2 [1 - \cos(2\bar{k}u_i + 2\bar{\phi})] \\ &= \frac{12}{N^3 \Delta u^2} \sum_{i=1}^N u_i^2 [1 - \cos(2\bar{k}u_i) \cos(2\bar{\phi})] \\ &= 1 + O\left(\frac{1}{N}\right), \end{aligned}$$

$$\begin{aligned} \sum_{i=1}^N v_1^i v_2^i &= \frac{\sqrt{2}}{N} \sum_{i=1}^N \cos(\bar{k}u_i + \bar{\phi}) = \frac{\sqrt{2}}{N} \sum_{i=1}^N \cos(\bar{k}u_i) \cos(\bar{\phi}) \\ &= O\left(\frac{1}{N}\right), \end{aligned}$$

$$\begin{aligned} \sum_{i=1}^N v_1^i v_3^i &= \frac{\sqrt{2}}{N} \sum_{i=1}^N \sin(\bar{k}u_i + \bar{\phi}) = \frac{\sqrt{2}}{N} \sum_{i=1}^N \cos(\bar{k}u_i) \sin(\bar{\phi}) \\ &= O\left(\frac{1}{N}\right), \end{aligned}$$

$$\sum_{i=1}^N v_2^i v_3^i = \frac{1}{N} \sum_{i=1}^N \sin(2\bar{k}u_i + 2\bar{\phi}) = O\left(\frac{1}{N}\right),$$

$$\begin{aligned}
\sum_{i=1}^N v_1^i v_4^i &= \frac{24}{N^2 \Delta u} \sum_{i=1}^N u_i \sin(\bar{k}u_i + \bar{\phi}) = O\left(\frac{1}{N}\right), \\
\sum_{i=1}^N v_2^i v_4^i &= \frac{12}{N^2 \Delta u} \sum_{i=1}^N u_i \sin(2\bar{k}u_i + 2\bar{\phi}) = O\left(\frac{1}{N}\right), \\
\sum_{i=1}^N v_3^i v_4^i &= \frac{12}{N^2 \Delta u} \sum_{i=1}^N u_i [1 - \cos(2\bar{k}u_i + 2\bar{\phi})] \\
&= \frac{12}{N^2 \Delta u} \sum_{i=1}^N u_i \sin(2\bar{k}u_i) \sin(2\bar{\phi}) = O\left(\frac{1}{N}\right). \quad (C5)
\end{aligned}$$

These inner products show that the v_i s are approximately orthonormal. In other words, V is nearly orthogonal, $V^T V = I_{N \times N} + E$ with $|E| = O(1/N)$.

Therefore, the information matrix:

$$D^T D \approx \Lambda^2 = \begin{bmatrix} N & 0 & 0 & 0 \\ 0 & \frac{N}{2} & 0 & 0 \\ 0 & 0 & \frac{N}{2} & 0 \\ 0 & 0 & 0 & \frac{N^3 \Delta u^2}{24} \end{bmatrix}. \quad (C6)$$

Using expression $D^\dagger = (D^T D)^{-1} D^T$ for the pseudoinverse yields

$$D^\dagger \approx \Lambda^{-1} V^T. \quad (C7)$$

As an immediate application of Eq. (C7), we compute the systematic errors due to ignoring the fringe envelope expressed in Eq. (47). Using Eq. (C7) enables to write

$$\delta X_i^{\text{env}} = (D^\dagger \delta y^{\text{env}})_i \approx \Lambda_{ii}^{-1} v_i^T \delta y^{\text{env}}, \quad i = 1, 2, 3, 4. \quad (C8)$$

Inserting definitions in Eq. (C1) and performing the same kind of computations as Eq. (C5) yield

$$\begin{aligned}
\delta X^{\text{env}} &\approx \mu_2 \bar{I} \bar{V} \begin{bmatrix} -\frac{1}{2N} \sum_i u_i^2 \cos(\bar{k}u_i + \bar{\phi}) \\ -\frac{1}{N} \sum_i u_i^2 \cos^2(\bar{k}u_i + \bar{\phi}) \\ \frac{1}{2N} \sum_i u_i^2 \cos(\bar{k}u_i + \bar{\phi}) \sin(\bar{k}u_i + \bar{\phi}) \\ \frac{12}{N^3 \Delta u^2} \sum_i u_i^3 \cos(\bar{k}u_i + \bar{\phi}) \sin(\bar{k}u_i + \bar{\phi}) \end{bmatrix} \\
&\approx \mu_2 \bar{I} \bar{V} \begin{bmatrix} O(N) \Delta u^2 \\ O(N^2) \Delta u^2 \\ O(N) \Delta u^2 \\ O(1) \Delta u \end{bmatrix}, \quad (C9)
\end{aligned}$$

where the delay $a_1 + d$ in the envelope expression (47) has been neglected compared with u_i . Using relation (33), we get

$$\begin{aligned}
\delta I^{\text{env}} &= O(N \mu_2 \Delta u^2 \bar{I} \bar{V}), & \delta (IV)^{\text{env}} &= O(N^2 \mu_2 \Delta u^2 \bar{I} \bar{V}), \\
\delta \phi^{\text{env}} &= O(N \mu_2 \Delta u^2), & \delta k^{\text{env}} &= O(\mu_2 \Delta u). \quad (C10)
\end{aligned}$$

In view of Eq. (C10), only the wavenumber calibration error due to neglecting the fringe envelope does not grow as the stroke length increases. The intensity and phase calibration errors grow linearly with the stroke length while the visibility calibration error grows quadratically.

Appendix D: Information Matrix

We now turn to study the information matrix for using the envelope projection technique described in Subsection 3.C, which is $D^T P D$. Let us first find P . To do this, we define normalized vectors,

$$\tilde{e}_1 = \sqrt{\frac{24}{N^3 \Delta u^2}} e_1, \quad \tilde{e}_2 = \sqrt{\frac{160}{N^5 \Delta u^4}} e_2, \quad \tilde{e}_3 = \sqrt{\frac{160}{N^5 \Delta u^4}} e_3. \quad (C11)$$

Using the asymptotic behaviors in Eq. (A6) of C_i and S_i , it is straightforward to verify that, to the leading order, \tilde{e}_i , $i = 1, 2, 3$ are orthonormal. Therefore, $P \approx 1 - \sum_{i=1}^3 \tilde{e}_i \tilde{e}_i^T$. The information matrix for using envelope projection is

$$D^T P D = \Lambda V^T P V \Lambda \approx \Lambda^2 - \sum_{i=1}^3 \Lambda V^T \tilde{e}_i \tilde{e}_i^T V \Lambda. \quad (C12)$$

The matrix multiplications $V^T \tilde{e}_i$, $i = 1, 2, 3$, can be carried out in the same fashion as in Eq. (C5) by using the asymptotic behaviors in Eq. (A6) of C_i , S_i , and the sums in Eq. (A11). After tedious calculations, to the leading order, we get

$$\begin{aligned}
V^T \tilde{e}_1 &\approx 0, & V^T \tilde{e}_2 &\approx (0, \sqrt{5}/3, 0, 0)^T, \\
V^T \tilde{e}_3 &\approx (0, 0, \sqrt{5}/3, 0)^T. \quad (C13)
\end{aligned}$$

Inserting the results from Eq. (C13) into Eq. (C12) gives

$$\begin{aligned}
D^T P D &\approx \Lambda \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 4/9 & 0 & 0 \\ 0 & 0 & 4/9 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \Lambda \\
&= \begin{bmatrix} N & 0 & 0 & 0 \\ 0 & \frac{2N}{9} & 0 & 0 \\ 0 & 0 & \frac{2N}{9} & 0 \\ 0 & 0 & 0 & \frac{N^3 \Delta u^2}{24} \end{bmatrix}. \quad (C14)
\end{aligned}$$

The asymptotic expressions (C6) and (C14) show that the difference between $D^T P D$ and $D^T D$ is a reduction

factor 4/9 in the second and third diagonal values. Thus using the envelope projection method causes the variances of the visibility and dispersion phase to increase by a factor 9/4. The variances of the wave-number and intensity remain unchanged.

Appendix E: Proof of Eq. (43) in Subsection 3.C

We will show that, for each $i = 1, \dots, 4$,

$$E_i \equiv v_i^T \text{diag}(\cos(\bar{k}u_1 + \bar{\phi}), \dots, \cos(\bar{k}u_N + \bar{\phi}))v_i = O(1/N), \quad (\text{E1})$$

by computing the sum directly. Using the definitions of v_i s in Eq. (C1) and some basic trigonometric relations we get

$$\begin{aligned} E_1 &= \frac{1}{N} \sum_{i=1}^N \cos(\bar{k}u_i + \bar{\phi}) = \frac{1}{N} \sum_{i=1}^N \cos(\bar{k}u_i) \cos \bar{\phi} = O(1/N), \\ E_2 &= \frac{2}{N} \sum_{i=1}^N \cos(\bar{k}u_i + \bar{\phi}) \cos^2(\bar{k}u_i + \bar{\phi}) \\ &= \frac{1}{N} \sum_{i=1}^N [1 + \cos(2\bar{k}u_i + 2\bar{\phi})] \cos(\bar{k}u_i + \bar{\phi}) \\ &= \frac{1}{2N} \sum_{i=1}^N [3 \cos(\bar{k}u_i + \bar{\phi}) + \cos(3\bar{k}u_i + 3\bar{\phi})] = O(1/N), \\ E_3 &= \frac{2}{N} \sum_{i=1}^N \cos(\bar{k}u_i + \bar{\phi}) \sin^2(\bar{k}u_i + \bar{\phi}) = E_1 - E_2 = O(1/N), \\ E_4 &= \frac{24}{N^3 \Delta u^2} \sum_{i=1}^N u_i^2 \cos(\bar{k}u_i + \bar{\phi}) \sin^2(\bar{k}u_i + \bar{\phi}) \\ &= \frac{6}{N^3 \Delta u^2} \sum_{i=1}^N u_i^2 \sum_{i=1}^N u_i^2 [\cos(\bar{k}u_i + \bar{\phi}) - \cos(3\bar{k}u_i + 3\bar{\phi})] \\ &= \frac{6}{N^3 \Delta u^2} \sum_{i=1}^N u_i^2 [\cos(\bar{k}u_i) \cos \bar{\phi} - \cos(3\bar{k}u_i) \cos(3\bar{\phi})] \\ &= O(1/N) \end{aligned} \quad (\text{E2})$$

where we have used the asymptotic behavior (A6) of $C_l(k, \Delta u)$ for $l = 0, 2$.

We thank Mike Heflin and Lisa Sievers for reviewing the manuscript. Many helpful discussions with Mike Heflin, Janice Shen, and Jeff Yu are appreciated. This work was prepared at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

References

1. J. Marr, "Space Interferometry Mission (SIM) overview and current status," *Proc. SPIE* **4852**, 1–15 (2003).
2. R. A. Laskin, "SIM technology development overview-light at the end of the tunnel," *Proc. SPIE* **4852**, 16–32 (2003).
3. R. Goullioud and T. J. Shen, "MAM testbed detailed description and alignment," in *Proceedings of IEEE Conference on Aerospace* (IEEE, 2004), pp. 2191–2200.
4. R. Goullioud and T. J. Shen, "SIM astrometric demonstration at 24 picometers on the MAM testbed," in *Proceedings of IEEE Conference on Aerospace* (IEEE, 2004), pp. 2179–2190.
5. M. Milman, C. Zhai, and M. Regehr, "White-light interferometry using a channeled spectrum. 1. General models and fringe estimation algorithms," *Appl. Opt.* **46**, 5853–5865 (2007).
6. X. Pan, F. Zhao, and M. Shao, "MAM testbed data analysis: cyclic averaging," *Proc. SPIE* **4852**, 612–622 (2003).
7. R. C. M. Learner, A. P. Thorne, I. Wynnee-Jones, J. W. Brault, and M. C. Abrams, "Phase correction of emission line Fourier transform spectra," *J. Opt. Soc. Am. A* **12**, 2165–2171 (1995).
8. K. Rahmelow and W. Hubner, "Phase correction in Fourier transform spectroscopy," *Appl. Opt.* **36**, 6678–6686 (1997).
9. C. D. Barnett, J. M. Blaisdell, and J. Susskind, "Practical method for rapid and accurate computation of interferometric spectra for remote sensing application," *IEEE Trans. Geosci. Remote Sens.* **38**, 169–183 (2000).
10. A. Ben-David and A. Ifarraguerri, "Computation of a spectrum from a single-beam Fourier-transform infrared interferogram," *Appl. Opt.* **41**, 1181–1189 (2002).
11. M. Regehr and M. Milman, "Analysis and numerical modeling of error sources in SIM star light phase detection," in *Proceedings of IEEE Conference on Aerospace* (IEEE, 2005), pp. 1–8.
12. M. Milman, "Accurately computing the optical pathlength difference for a Michelson interferometer with minimal knowledge of the source spectrum," *J. Opt. Soc. Am. A* **22**, 2774–2785 (2005).
13. M. Born and E. Wolf, *Principles of Optics*, 6th ed. (Pergamon, 1989).
14. P. R. Lawson, "Phase and group delay estimation," in *Principles of Long Baseline Stellar Interferometry*, P. R. Lawson, ed., Course Notes from the 1999 Michelson Interferometry Summer School (Jet Propulsion Laboratory, Calif., 1999).
15. D. G. Luenberger, *Optimization By Vector Space Methods* (Wiley, 1969).
16. G. Golub and V. Pereyra, "The differentiation of pseudoinverse and nonlinear least squares problems whose variable separate," *SIAM (Soc. Ind. Appl. Math.) J. Numer. Anal.* **10**, 413–432 (1973).
17. R. Fletcher, *Practical Methods for Optimization* (Wiley, 1987).
18. M. Milman and S. G. Turyshev, "Observational model for microarcsecond astrometry with the Space Interferometer Mission," *Opt. Eng.* **42**, 1873–1883 (2003).
19. A. J. Pickles, "A stellar spectral flux library: 1150–25000 Å," *Publ. Astron. Soc. Pac.* **110**, 863–878 (1998).